INS
C1

Wound Healing and Orofacial Clefting

The present invention relates to the isolation of a nucleic acid molecule and the protein encoded thereby; and the use of these products as therapeutic agents particularly, but not exclusively, in gene therapy and/or tissue repair such as, without limitation enhancing wound and tissue healing and for the treatment of orofacial clefting.

Background to the Invention

Orofacial clefting is the most common human malformation, with an overall incidence of approximately 1 in 600 births. Cleft palate (CP) requires complex management and follow up by multi-disciplinary medical and surgical teams. It has a major impact on both communication and psychological morbidity. Both animal and human studies have shown that cleft palate can be the end result of a number of different aetiological processes. Amongst the teratogenic agents that can cause CP are common drugs (diazepam, sodium valproate, alcohol etc.). Maternal diabetes also confers an increased risk of CP. However, twin studies and familial segregation analyses in a number of different populations have consistently shown that there is a major genetic component to the aetiology of this common developmental abnormality. Despite this, little is known about the specific genetic defects underlying CP. The care of children with clefting is now concentrated in supra-regional centres where the necessary resources can be assembled.

Isolated cleft palate (CPO) is a common human malformation, with a total birth incidence of 1 in 1250 in the West of Scotland [FitzPatrick et al. 1994]. Significant familial clustering [Carter et al. 1982; Shields et al., 1981; Christensen & Fogh-Andersen 1993] and twin studies [Shields et al. 1979] have both suggested that there is a major genetic component in the etiology of CPO. These studies and others [Fogh-Andersen, 1942] have also shown that CPO and cleft lip with or without cleft palate (CL(P)) are genetically distinct subgroups of orofacial clefting. CPO is a

common feature of chromosomal abnormalities, affecting ~15% of all cases of simple autosomal aneuploidy (Schinzel 1994) and is associated with more than 370 different malformation syndromes (Baraitser & Winter, 1997). However, in non-syndromic CL(P) or CPO, relative risk ratio analyses have indicated that there may be a relatively small number of interacting causative loci [FitzPatrick & Farrell, 1993; Christiensson & Mitchell, 1996]. As yet, no disease-causing mutations in non-syndromic CPO have been identified. In the present application we identify for the first time a previously unrecognised gene for cleft palate located at 2q32.

Development of the secondary palate, particularly in mice, has been used as a paradigm in developmental biology, and extensive descriptive studies of the distribution of proteins with putative roles in this process have been published. This knowledge base has led to the extensive use of candidate gene approaches in attempts to unravel the genetic basis of human cleft palate. However, these studies have met with limited success. Purely genetic approaches, in contrast, by the identification of genes causing rare syndromic forms of cleft palate, have yielded greater insights. An example of this is the positional cloning of the gene responsible for Treacher Collins syndrome. Experimental mouse models are another source of valuable information on clefting, for example the transgenic knock outs of Tgfβ3 and Msx1. However, extrapolation from these models by testing for genetic association in human populations has been generally unsuccessful.

There are a number of semi-dominant and recessive mouse models of isolated cleft palate, for example Twirler (Tw) and Dancer (Dc). These provide excellent opportunities to identify potential susceptibility loci for human CP via the positional cloning of the mouse genes. However, none of the loci for these mouse models lies within a region of conserved synteny with the human chromosome 2q32 region a gene from which is the subject of this application.

The problem of identifying susceptibility loci for human malformations has recently been addressed by utilising the common phenomenon of autosomal aneuploidy.

Using the Human Cytogenetics Database, the entire set of information available on post-natally ascertained cases of simple autosomal aneuploidy has been statistically analysed in a manner not previously attempted. This has allowed the identification of specific autosomal regions that are significantly associated with particular malformations. 37 different malformations have been studied, covering a variety of different developmental processes. For cleft palate, 5 different putative loci were identified. The validity of the approach was confirmed by the identification of loci on 4p and 4q, which have previously been suggested to harbour CP susceptibility genes as a result of association or linkage studies. Unexpectedly a major new locus has been identified at 2q32-q33. Interestingly, the penetrance of haploinsufficiency in causing cleft palate at this locus appears to be higher than that of the other four loci. This region has not previously been implicated in the pathogenesis of CP. We have therefore gone on to obtain further independent evidence for the importance of this locus.

From our studies we have identified a human gene which, when dysfunctional, unexpectedly results in cleft palate in man. Replacement of gene function is therefore useful in the treatment of cleft palate and the related cleft lip. There is precedent that disruption of the TGFβ3 gene, for example, leads to the development of cleft palate in animal models. There is also precedent that the use of TGFβ3, for example, is beneficial in the treatment of wounds to achieve enhanced rates of healing. The gene and the protein of the present invention are also therefore useful therapeutically to enhance wound healing. They are also useful clinically to improve tissue repair or regeneration in other clinical contexts, for example in inducing repair of damage to cartilage or boney tissue.

In the present application, we believe we have demonstrated the unexpected existence of an important locus on human chromosome 2q32 causing cleft palate. We have cloned this CP gene and found mutations in a large cohort of patients with cleft palate. Given the problems of day-to-day clinical management of patients with this distressing condition, we also expect that the diagnostic tools described herein

will be rapidly exploited clinically and possibly suggest approaches to prevention strategies for this common malformation.

It is therefore an object of the present invention to provide a tissue repair gene and/or protein for use as a therapeutic agent.

It is a further object of the present invention to provide a tissue repair gene and/or protein for use as a diagnostic agent.

Statements of the Invention

According to a first aspect of the present invention is an isolated nucleic acid encoding a tissue repair protein, the nucleic acid may be selected from the group consisting of:

(a) DNA having the nucleotide sequence given herein as SEQ ID NO:1 (which encodes the protein having the amino acid sequence given herein as SEQ ID NO:2), and which encodes a tissue repair protein;

(b) nucleic acids which hybridize to DNA of (a) above (e.g., under stringent conditions) and which encode a tissue repair protein ; and

(c) nucleic acids which differ from the DNA of (a) or (b) above due to the degeneracy of the genetic code, and which encode a tissue repair protein encoded by a DNA of (a) or (b) above.

DNAs of the present invention include those coding for proteins homologous to, and having essentially the same biological properties as, the proteins disclosed herein, and particularly the DNA disclosed herein as SEQ ID NO:1 and encoding the protein given herein SEQ ID NO:2 This definition is intended to encompass natural allelic variations therein. Thus, isolated DNA or cloned genes of the present invention can be of any species of origin, including mouse, rat, rabbit, cat, porcine, and human, but are preferably of mammalian origin. For example, the mouse homologue SEQ ID NO:3 of the human protein encoded by SEQ ID NO:1, differs by only 2 amino acids

in a total of 733 amino acids. Thus, DNAs which hybridize to DNA disclosed herein as SEQ ID NO:1 (or fragments or derivatives thereof which serve as hybridization probes as discussed below) and which code on expression for a protein of the present invention (e.g., a protein according to SEQ ID NO:2), i.e. the tissue repair protein of the present invention associated with orofacial clefting and/or wound healing are to be included in the definition.

Conditions which will permit other DNAs which code on expression for a protein of the present invention to hybridize to the DNA of SEQ ID NO:1 disclosed herein can be determined in accordance with known techniques. For example, hybridization of such sequences may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions (e.g., conditions represented by a wash stringency of 35-40% Formamide with 5x Denhardt's solution, 0.5% SDS and 1x SSPE at 37°C; conditions represented by a wash stringency of 40-45% Formamide with 5x Denhardt's solution, 0.5% SDS, and 1x SSPE at 42°C; and conditions represented by a wash stringency of 50% Formamide with 5x Denhardt's solution, 0.5% SDS and 1x SSPE at 42°C, respectively) to DNA of SEQ ID NO:1 disclosed herein in a standard hybridization assay. See, e.g., J. Sambrook et al., *Molecular Cloning, A Laboratory Manual* (2d Ed. 1989) (Cold Spring Harbor Laboratory). In general, sequences which code for proteins of the present invention and which hybridize to the DNA of SEQ ID NO:1 disclosed herein will be at least 75% homologous, 85% homologous, and even 95% homologous or more with SEQ ID NO:1. Further, DNAs which code for proteins of the present invention, or DNAs which hybridize to that given as SEQ ID NO:1, but which differ in codon sequence from SEQ ID NO:1 due to the degeneracy of the genetic code, are also an aspect of this invention. The degeneracy of the genetic code, which allows different nucleic acid sequences to code for the same protein or peptide, is well known in the literature. See, e.g., U.S. Patent No. 4,757,006 to Toole et al. at Col. 2, Table 1.

According to a further aspect of the present invention is an isolated nucleic acid encoding a tissue repair protein, the nucleic acid may be selected from the group consisting of:

(a) DNA having the nucleotide sequence given herein as SEQ ID NO:3 (which encodes the protein having the amino acid sequence given herein as SEQ ID NO:4), and which encodes a tissue repair protein;

(d) nucleic acids which hybridize to DNA of (a) above (e.g., under stringent conditions) and which encode a tissue repair protein ; and

(e) nucleic acids which differ from the DNA of (a) or (b) above due to the degeneracy of the genetic code, and which encode a tissue repair protein encoded by a DNA of (a) or (b) above.

According a further aspect of the invention there is therefore provided a cloned nucleic acid molecule encoding a tissue repair protein contained in a Yeast Artificial Chromosome species designated as AB 1380 YAC-CP1 and deposited with NCIMB Limited of Aberdeen, Scotland (UK) under accession number NCIMB 41005.

According to a yet further aspect of the invention there is provided a nucleic acid molecule which encodes a tissue repair protein and comprises a nucleotide sequence which hybridises to the nucleic acid of SEQ ID NO:1 under high stringency conditions.

Preferably, hybridisation occurs under stringent conditions such as 1 x SSC, 0.1% SDS at 65 °C.

Preferably, the nucleic acid is mammalian in origin and more preferably is human.

According to a yet further aspect of the invention there is provided use of the nucleic acid of SEQ ID NO:1, fragments and/or variants thereof, in determining expression of mRNA in selected target tissue(s) for diagnosing cleft palate.

According to a yet further aspect of the invention there is provided use of the nucleic acid of SEQ ID NO:1, fragments and/or variants thereof, in determining the presence of mutants in the DNA and thus diagnosing patients suffering from cleft palate.

According to a further aspect of the invention there is provided a polypeptide, or a protein comprising an epitope for an antibody or a protein modified by one or more amino acid modifications and comprising an epitope, or a fragment modified or unmodified comprising an eptitope for a tissue repair protein encoded by SEQ ID NO:2. Ideally the polypeptide is encoded by the nucleic acid molecule of SEQ ID NO:1.

According to a yet further aspect of the invention there is provided a delivery vehicle comprising the isolated nucleic acid molecule or polypeptide of the invention.

Reference herein to the term delivery vehicle is intended to include any vector whether a viral vector or otherwise for example, without limitation, an adenovirus, a retrovirus, a herpesvirus, a plasmid, a phage, a phagemid or a liposome.

Ideally said delivery vehicle is adapted for administration, for example, but without limitation, by suitable formulation into a suspension.

More preferably, said delivery vehicle is adapted to deliver said nucleic acid molecule or polypeptide to selected tissue. Thus the delivery vehicle is provided with means to enable the nucleic acid molecule or polypeptide to be targeted to a specific site. The nature of the means comprises conventional technologies well known to those skilled in the art for example, without limitation, in the instance where the delivery vehicle is a viral vector said viral vector is provided with surface protein adapted to ensure the viral vector binds to and/or penetrates specific target tissues. Thus, in this way, the nucleic acid molecule or peptide, fragments or derivatives thereof of the invention can be used in gene therapy treatments.

According to a yet further aspect of the invention there is provided antibodies raised against the polypeptide, fragment or derivative thereof, of the invention. Ideally the antibodies are monoclonal and more ideally genetically engineered to be humanised. It will be apparent to those skilled in the art that the antibodies of the invention can be used to determine the expression of the polypeptide of the invention in selected target tissue and thus aid in the diagnosis of patients suffering from cleft palate.

According to a yet further aspect of the invention there is provided use of antibodies, fragments or derivatives thereof in diagnosis of orofacial clefting. It will be appreciated that the fragments or derivatives of the antibodies contain the epitope.

According to a yet further aspect of the invention there is provided a method for detecting the antibodies as described above, in a sample, comprising contacting with the sample immobilised antibody against a protein or protein fragment of SEQ ID NO:2, which antibody has bound thereto a labelled ligand comprising a protein or protein fragment of SEQ ID NO:2, and detecting labelled ligand bound to immobilised antibody or labelled ligand bound to antibody in the sample.

According to a yet further aspect of the invention there is provided a method for the treatment of orofacial clefting comprising administering to a patient suffering from orofacial clefting the nucleic acid molecule and/or polypeptide of SEQ ID NO:1 and/or SEQ ID NO:2.

According to a further aspect of the invention there is provided a method for the treatment of wounds comprising administering to a patient suffering from tissue damage the nucleic acid molecule and/or polypeptide of SEQ ID NO:1 and/or SEQ ID NO:2.

Preferably, the nucleic acid molecule and/or polypeptide is administered by the incorporation of said nucleic acid molecule into a delivery vehicle as herein described and ideally the method of treatment involves the use of gene therapy.

According to a yet further aspect of the invention there is the nucleic acid and/or protein,as herein before described for use as a pharmaceutical.

According to a yet further aspect of the invention there is provided use of the nucleic acid and/or protein, fragments or variants thereof of SEQ ID NO:1 and/or SEQ ID NO:2 for the manufacture of a medicament for the treatment of orofacial clefting and/or wound healing and/or tissue repair.

According to a yet further aspect of the invention there is provided an isolated nucleic acid molecule, fragment or variant thereof which encodes a tissue repair protein wherein said isolated nucleic acid molecule has a nucleotide sequence which hybridises to the nucleic acid of SEQ ID NO:3 under high stringency conditions.

Preferably, hybridisation occurs under stringent conditions such as 1 x SSC, 0.1% SDS at 65 °C.

Preferably, the nucleic acid is murine in origin.

According to a further aspect of the invention there is provided a polypeptide, or a protein comprising an epitope for an antibody or a protein modified by one or more amino acid modifications and comprising an epitope, or a fragment modified or unmodified comprising an eptitope for a tissue repair protein encoded by SEQ ID NO:4. Ideally the polypeptide is encoded by the nucleic acid molecule of SEQ ID NO:3.

According to a yet further aspect of the invention there is provided a delivery vehicle comprising the isolated nucleic acid molecule or polypeptide, fragments or derivatives thereof of SEQ ID NO:3 and/or SEQ ID NO:4.

According to a yet further aspect of the invention there is provided antibodies raised against the polypeptide, fragment or derivative thereof of SEQ ID NO:3. Ideally the antibodies are monoclonal.

According to a yet further aspect of the invention there is provided a method of producing a transgenic mammal comprising disrupting a gene, or the effective part thereof, the gene encoding at least one tissue repair protein.

Reference herein to disruption is intended to include complete or partial disruption of expression of the tissue repair protein such that the transgenic animal is unable to express levels of the said protein that are typically found in individuals suffering from cleft palate.

Preferably, the transgenic mammal is a rodent and ideally a mouse and more preferably the gene encoding the tissue repair protein is the nucleic acid molecule or fragment or derivative thereof of SEQ ID NO:3.

Preferably, the transgenic mammal is a human and more preferably the gene encoding the tissue repair protein is the nucleic acid molecule or fragment or derivative thereof of SEQ ID NO:1.

According to a yet further aspect of the invention there is provided a reporter gene construct based on the promoter region of a gene, or effective part thereof, encoded by SEQ ID NO:1.

According to a yet further aspect of the invention there is provided use of a reporter gene construct based on the promoter region of a gene, or effective part thereof, encoded by SEQ ID NO:1 in the detection/screening of pharmaceuticals and/or other compounds and their potential teratogenic effects.

Brief Description of the Figures

Embodiments of the invention will now be described by way of example only with reference to the following figures wherein:

Figures 1 A and 1 B represent chromosome painting of chromosome 2q in patients case 1 and case 2 respectively;

Figures 2 A and B represent fluorescent in situ hybridisation of a YAC probe encoding IGF binding protein 5 in patients case 1 and case 2 respectively;

Figures 3 A and B represent fluorescent in situ hybridisation of YAC's containing the markers D25311 and D2S309 respectively, in patients case 1;

Figure 4 represents fluorescent in situ hybridisation of YAC-CP1 in patient case 1;
Figure 5 represents fluorescent in situ hybridisation of YAC-CP1 in patient case 2;

Figure 6 represents the DNA nucleotide sequence of a human tissue repair gene SEQ ID NO:1;

Figure 7 represents the amino acid sequence for the tissue repair protein SEQ ID NO:2, encoded by SEQ ID NO:1;

Figure 8 represents the DNA nucleotide sequence of a mouse tissue repair gene SEQ ID NO:3;

Figure 9 represents the amino acid sequence for the tissue repair protein SEQ ID NO:4, encoded by SEQ ID NO:3;

Table 1 represents oligonucleotides used in the study; and

Table 2 represents the genetic map of the CP-1 region.

Detailed Description of the Invention

We have tested the novel hypothesis that there is a major locus for cleft palate located at human chromosome 2q32. We have cloned and characterised this cleft palate gene.

It is good clinical practice to perform karyotype analysis on all children with cleft palate. By this approach an 11 year old girl was ascertained with a *de novo* balanced reciprocal translocation between chromosomes 2q32 and 11p14. This patient has a midline posterior cleft of the soft palate, mild learning difficulties and subtle craniofacial dysmorphism (Case 2). A second patient was then studied (Case 1) with a similar clinical phenotype and another de novo balanced reciprocal translocation, this time between chromosomes 2q32 and 7p21. These patients have remarkably consistent clinical features. The existence of 2 different individuals with balanced translocations, possibly involving a common break point at 2q32, strongly supported the conclusion of aneuploidy studies, that an important susceptibility locus for cleft palate exists in this cytogenetic region.

We have therefore taken a molecular cytogenetic route to further analyse these two patients. Initially, we wished to establish that the two break points on 2q32 were indeed located in the same region. A number of experiments using single chromosome painting were performed, which confirmed that both re-arrangements were apparently simple balanced reciprocal translocations and that to a first approximation, they occurred at the same place on chromosome 2q (Figure 1). Next, we tested the possible involvement of a number of potential candidate genes in this region. Yeast artifical chromosomes (YACs) containing these candidates were isolated and each was used for fluorescent in situ hybridization (FISH) analysis. This showed, for example, that the fibronectin and IGF binding protein 5 (IGFBP5) genes are both telomeric to the breakpoints (Figure 2).

We then isolated a total of 70 YACs using genetic markers distributed across the whole 2q32-q33 region. FISH analysis was performed systematically using each of these YACs, in order to position the breakpoints more precisely. Typical results from this study are presented in Figure 3. In support of our initial hypothesis, the FISH analyses confirm that the 2q breakpoints in both patients do indeed lie in the same small interval. We narrowed this region to less than 2 centiMorgans, by virtue of the fact that YACs containing the marker D2S311 are centromeric to both patients' breakpoints, while YACs containing D2S309 are telomeric to the breakpoints, again in both patients (Figure 3 shows Case 1; the data for Case 2 are similar).

On the basis of this information, a YAC contig spanning this interval has been constructed. This has enabled us to resolve a number of uncertainties in current genetic maps of this region. We have placed 23 genetic markers within this small interval and isolated a total of 33 corresponding genomic clones. Our FISH results narrowed the region containing both breakpoints to <600 kb. The YAC clone CP-1 (Fig. 4, 5) crosses both breakpoints. Clearly, this means that we have cloned the novel CP gene. The novel gene is defined by the break points in patients "Case 1" (0213) and "Case 2" (0145) at the point where they occur within YAC-CP1 which has been deposited with the NCIMB Limited under accession number 41005. Cell lines from Case 1 are available in the form of lymphoblastoid cells from ECACC (Porton Down, UK).

The YAC clone spanning both breakpoints was used to screen flow-sorted, chromosome 2 specific cosmid and PAC libraries. A PAC and cosmid contig of the breakpoint region was constructed by a combination of fingerprinting and sequencing. Individual clones were used in further FISH analysis. This approach leads us to the PAC "CP-1" which has also been deposited as above which defines the region harbouring the CP gene. We then took the most direct route to isolating the gene itself, which was sequencing of this PAC clone.

The genomic DNA sequence was used for gene identification (SEQ ID NO:1) by a combination of EST database searching and computational prediction. We screened cDNA libraries to isolate clones corresponding to transcripts in this region. Together with the genomic sequence, these clones provided us directly with information about the intron-exon structure of the gene, and its promoter. In parallel studies, we also isolated the corresponding mouse gene (SEQ ID NO:3) by techniques well known in the art, which will be a prerequisite for transgenic studies.

Availability of the CP gene now allows us to analyse material from patients with cleft palate. We can look for microdeletions at this locus, using FISH analysis of patients with apparently normal karyotypes. We also undertake sequencing of the gene, using cDNA or genomic DNA substrates as appropriate, in these patients. These diagnostic tests are useful clinically in the management of patients with cleft palate and in genetic counselling of them and their families.

In view of evidence implicating a number of known biologically active molecules, both endogenous and exogenous, in the aetiology of CP, it is interesting to examine the influence of such agents on the expression of the newly isolated 2q32 CP gene. This involves the construction of reporter plasmids containing a reporter gene construct of the gene of SEQ ID NO:1 to examine promoter function of the CP gene. Such studies can then be followed up by a more detailed functional analysis of the promoter sequences including deletion mapping and by direct examination of the transcriptional response of the gene to a variety of agents by methods well known in the art. Thus the subject of the present invention is of value to the pharmaceutical industry for the toxicological evaluation of potential new drugs. Analysis of their effects on expression of the CP gene which is the subject of the present invention, will enable prediction of their possible teratogenicity.

As discussed above, using SEQ ID NO:3, we isolated mouse genomic clones from libraries of strain 129 DNA isogenic with the ES cells in use in our laboratories, by techniques known in the art, having confirmed the organisation of the mouse gene,

we performed transgenic knockout experiments using standard approaches. One specific approach involved the use of the IRES-βgeo targetting construct, which as well as eliminating function of the CP gene, allows monitoring of its expression pattern in the embryo by simple staining for β-galactosidase rather than extensive use of in situ hybridization. Such a transgenic approach also facilitated analysis of the effects of recognised teratogens on the expression of this gene in vivo.

Availability of a transgenic model enabled us to undertake other interesting studies. For example, as the mutation in mouse is variable in penetrance, we are in a position to utilise mouse genetic approaches to map modifier loci. These are likely to have homologues in man which could well be of clinical significance.

We have studied two unrelated individuals with strikingly similar clinical features, in whom there are apparently balanced de novo cytogenetic rearrangements involving the same region of chromosome 2q. We now describe molecular cytogenetic analyses that have localised the translocation breakpoint in both cases to a region of some 0.3 Mb between markers D2S311 and D2S115. This suggests that the true location of these breakpoints is 2q32. Independent support for the existence of a novel locus for cleft palate on 2q32 was obtained by a detailed statistical analysis on all cases in the Human Cytogenetics Database of non-mosaic single contiguous autosomal deletions associated with orofacial clefting. This revealed 2q32 to be one of only three chromosomal regions in which haploinsufficiency is highly significantly associated with isolated cleft palate. In combination, our data provided strong evidence for the existence at 2q32 of a gene that is critical to the development of the secondary palate. The close proximity of the two chromosomal breakpoints also made the positional cloning of this gene a realistic possibility.

Subjects and Methods

Case 1: Case 1, the fourth child of healthy, non-consanguineous parents, was delivered at 38 weeks' gestation, weighing 2.95kg. Cleft palate was noted at birth. Delayed motor development was apparent at four months and the patient did not

walk until two years. Particular problems were noted with the acquisition of language skills. She underwent repair of her cleft palate at eighteen months of age and required pharyngoplasty at eleven years. Her hearing was normal. She had a prominent nasal bridge, a small mouth and long, slender fingers. Her growth has been satisfactory, and her height has always been on or above the 50th centile. Her weight was below the 10th centile until the age of five but at the age of ten was on the 75th centile. Her head circumference was on the 50th centile. She has moderate learning disability. Blood chromosome analysis revealed an apparently balanced reciprocal translocation with the karyotype 46XX, t(2;7)(q33;p21). Parental karyotypes were normal. FISH analysis showed no deletion of 22q11.22.

Case 2: Case 2 was a female delivered at term after an uneventful pregnancy. At birth she was noted to have cleft palate and minor facial dysmorphism. In addition to repair of her palate, she had required surgical correction of a convergent squint. On examination at the age of 8 years, she has fair hair and skin, a long, narrow face with apparent hypotelorism, a prominent nasal bridge and pinched appearance above the nares, and a small mouth and jaw. She had abnormal dermatoglyphics with a reduced ridge count. She was of slender build, with height between the 75th and 90th centiles and weight on the 10th centile. She had mild global developmental delay, particularly in language skills, and was one year behind her peers in a mainstream school. Chromosome analysis revealed an apparently balanced reciprocal translocation with the karyotype 46,XX,t(2;11)(q33;p14). Parental karyotypes were normal. FISH analysis showed no deletion of 22q11.22.

Molecular cytogenetic analysis

Our initial working map across the 2q breakpoint region (Dib et al. 1996) and a modified physical map, based on the consensus map of Collins et al. (1996), adjusted to reflect data that have emerged from our FISH studies, are presented (Table 2). In addition to YACs containing this set of genetic markers, we also isolated YAC clones containing a number of genes that were known to be located in this region but

had not been finely mapped. These were selected on the basis of their potential involvement in the etiology of cleft palate (Table I).

Statistical analysis of chromosomal deletions

By analysing all cases of single, contiguous, non-mosaic autosomal deletions stored in the Human Cytogenetics Database (HCDB) (Schinzel, 1994) three chromosomal regions (2q32, 4p16-13, 4q31-35) were identified where monosomy is non-randomly associated with CPO. However, HCDB searches alone do not differentiate cleft palate in the context of CL(P) cases from those cases that have CPO. As these are etiologically distinct subgroups of orofacial clefting and might be expected to have different causative genetic loci, the precise significance of these three chromosomal loci must necessarily be unclear. For a better insight into the phenotypes associated with deletion of these regions, the original case reports of all cases were obtained and reviewed. A statistical reanalysis was then performed as before but confining the analysis to confirmed CPO cases. Briefly, the distribution of deletions of regions including bands on chromosomes 4 or 2q in CPO patients was determined. The observed number of CPO-associated deletions of each band was compared with the expected number, calculated from the distribution of all band deletions on chromosomes 4 and 2q in HCDB. The number of deletions of any band was taken to follow a Poisson distribution, since this number is usually small. Confidence limits for the observed number of deletions and the significance of any deviation from expectation were calculated as described by Vasarhelyi and Friedman (1989).

Results

Patient phenotypes

Initial clues to the existence of a CPO locus on 2q32-33 came through the identification of a patient (Case 2) with a de novo balanced reciprocal translocation t(2;11)(q33;p14). A second patient (Case 1) was identified, again with CPO and a de

novo translocation [t(2;7)(q33;p21)] involving the same cytogenetic band on chromosome 2. Both patients had strikingly similar clinical appearances.

FISH Analysis

Initial chromosome painting studies confirmed that the sizes of the translocated 2q fragments were approximately the same in both patients (Fig. 1A, 1B). To attempt to establish whether the breakpoints in Cases 1 and 2 had occurred within the same region of chromosome 2q32, single locus FISH analysis was then conducted, using a large collection of YACs containing markers mapping within the 2q32-33 region. The results of this FISH study are summarised in Table 1. YACs containing the candidate genes FN1 (fibronectin) (Iamaroon 1996), IGFBP5 (Ferguson et al. 1992), IGFBP2 (not shown) and IHH (Indian Hedgehog) (Leek et al 1997, not shown) were all found to map distal (telomeric) to both chromosome 2 breakpoints.

Markers flanking both patients' 2q breakpoints were next identified by FISH (Fig. 3A, B). The signal generated by YACs containing D2S311 is present on both the normal and derivative copies of chromosome 2, in both Cases 1 and 2. Thus D2S311 is proximal (centromeric) to both breakpoints; this is clearly shown through simultaneous hybridization to the centromeric probes D2Z1 and either D7Z1 or D11Z1. In contrast, FISH with YACs containing D2S309, or D2S116 gave signals lying distal to the breakpoint in both individuals. This suggested that both breakpoints lie within a common region of some 2.5 Mb of 2q32. To reduce this interval further, YACs 14HA2 (containing D2S2189), 26IF5 (containing D2S1384/D2S307/ CTLA4/D2S105/D2S72) and 11GG8 (containing D2S115) were analysed. All were found to map distal to the 2q32 breakpoint in both patients. In this way, by systematic FISH analysis, the breakpoint region in both these patients with cleft palate has been localized to an interval, which may be as small as 0.3 Mb, according to current maps (Collins et al. 1996).

We suggest that a previously unrecognised locus causing cleft palate resides in chromosome region 2q32. Our suggestion relies on the integration of clinical, cytogenetic, molecular and statistical data. Two unrelated children had strikingly similar clinical features, each having a de novo cytogenetic rearrangement apparently involving the same band on chromosome 2q. Both girls had cleft palate. They also both had mild learning difficulties and a strikingly similar facial appearance. While their facial dysmorphisms are subtle, these and other clinical features were reminiscent of those seen in velocardiofacial syndrome (VCFS, OMIM 192430). It would appear, however, that these girls do not have VCFS, as neither of them has a cardiac malformation, nor the microdeletion of 22q11.22 seen in most cases of VCFS (Scrambler et al., 1992). The existence of non-22q-deleted phenocopies of VCFS is well recognized (Daw et al., 1996), so that it is possible that a 2q32 locus accounts for a proportion of such cases.

We isolated markers for this region of chromosome 2q, which allowed us to isolate a large number of genomic reagents for precise delineation of the 2q breakpoints in both cases. The 30 separate YACs that have been isolated provide good representation of this 2.5 Mb region.

The initial G-banded cytogenetic studies performed on both patients indicated the existence of a common breakpoint near 2q33. However, our single locus FISH studies collectively strongly suggest that the breakpoint in both Case 1 and Case 2 lies in 2q32. This discrepancy may result from the known bias in reporting of breakpoints in favor of Giemsa-pale bands, or it may be that the true breakpoint is at the 2q32-2q33 band junction. More importantly, both breakpoints mapped to a very small chromosomal region between D2S311 and D2S115, strongly suggesting that the same single gene is disrupted in both patients. These flanking markers are estimated to map at 207.169 Mb and 207.462 Mb respectively, in the current version of the Location Database LDB (Collins et al. 1996; ttp://cedar.genetics. soton.ac.uk/public html/). The cleft palate gene was thus localized by our studies to a region as small as 0.3 Mb. Two additional genetic markers (D2S374 and D2S1413)

have been mapped into this interval. A genomic clone, YAC-CP1, crosses both patients' breakpoints and represents the cloning of this cleft palate gene. Proof that 2q32 contains a CPO-causative genetic locus has been provided by the cloning of a gene whose function is disrupted by both breakpoints in these patients and the demonstration of mutations within this gene in cytogenetically normal individuals with CPO.

REFERENCES

Beiraghi S, Foroud T, Diouhy S, Bixler D, Conneally PM, Delozierblanchet D and Hodes MS (1994). Clinical Genet., 46, 255-256.

Bitgood MJ, McMahon AP. (1995). Dev. Biol.;172:126-38.

Carter CO, Evans K, Coffey R, Roberts JAF, Buck A, Roberts MF (1982). J Med Genet. 19: 329-331.

Christensen K, Fogh-Andersen P. Cleft Palate-Craniofacial Journal 30:469-474.

Christensen K, Mitchell LE (1996). Am. J. Hum. Genet. 58: 182-190.

Collins A, Frezal J, Teague J Morton NE (1996). Nature Genet 13 458.

Dib C, Fauré S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J and Weissenbach J (1996). Nature, 380, 152-154.

Dixon MJ and Ferguson MWJ (1992) Arch. Oral Biol., 37, 395-410.

Ferguson MWJ (1988). Development (Suppl.), 103, 41-60.

Ferguson MWJ (1994). Nature Genet., 6, 329-330.

Ferguson MWJ, Sharpe PM, Thomas BL and Beck F (1992). J. Anat., 181, 219-238.

FitzPatrick DR, Kondaiah P, Dehnex F and Akhurst R (1990) Development, 109, 585-595.

FitzPatrick DR, Farrali M (1993). J. Craniofac. Genet. Dev. Biol.;13, 230-235.

Fitzpatrick DR, Raine PAM, Boorman JG (1994). J. Med. Genet.;31,126-129.

Fogh-Andersen P (1942). Inheritance of harelip and cleft palate: contribution to the elucidation of the etiology of the congenital clefts of the face. Copenhagen,.

Gashler AL, Bonthron DT, Madden SL, Rauscher FJ, Collins T and Sukhatme VP (1992). Proc. Natl. Acad. Sci, USA, 89, 10984-10988.

Hill L, Murphy M, McDowall M and Paul AH (1988). Journal of Epidemiology and Community Health, 42, 1-7.

Hogan B, Beddington R, Constantini F and Lacy E (1994). Manipulating the mouse embryo; A laboratory manual. Cold Spring Harbor Press.

Gingrich JC, Boehrer DM, Garnes JA, Johnson W, Wong BS, Bergmann A, Eveleth GG, Langiois RG and Carrano AV (1996). Genomics, 32, 65-74.

Iamaroon A and Diewert VM (1996). J. Craniofacial Genet. Dev. Biol., 16, 48-51.

Kaartinen V, Voncken JW, Shuler C, Warburton D, Bu D, Heirkerkamp N and Groffen J (1995). Nature Genet., 11, 415-421.

Leek JP, Moynihan TP, Anwar R, Bonthron DT, Markham AF, Lench NJ (1997). Cytogenet Cell Genet, 76,187-8.

Lidral AC, Romitti PA, Basart AM, Doetschman T, Leysens NJ, DaackHirsh S, Semina EV, Johnson LR, Machida J, Burds A, Parnell TJ, Rubenstein JLR and Murray JC (1998). Amer. J. Hum. Genet., 63, 557-568.

Liu HC, Mele C, Catz D, Noyes N, Rosenwaks Z (1995). J Assist Reprod Genet, 12:78-87.

Lyon MF, in : Genetic Variants and Strains of the Laboratory Mouse (1996). Ed. MF Lyon, S Rastan and SDM Brown. Vol. 1, third edition, Oxford University Press, Oxford.

Markus T and Booth P (1995). British Medical Journal, 311, 765.

Morishita M, Shiba R, Chiyo HA, Furuyama JI, Fujita H and Atsumi Y (1983). J. Oral. Maxillofacial Surgery, 41, 601-605.

Murray JC (1995) Am. J. Hum. Genet., 57, 227-232.

O'Rahilly R and Muller F (1987). Developmental stages in human embryos. Washington: Carnegie Institute of Washington, 204-208.

Ohsaki Y, Nagata K, Kurisu K (1995). Acta Anat (Basel), 153, 161-7.

Proetzel G, Pawlowski SA, Wiles MV, Yin MY, Boivin GP, Howles PN and Ding JX, Ferguson MWJ, Doetschman T (1995). Nature Genet., 11, 409-414.

Resnick N, Collins T, Atkinson W, Bonthron DT, Dewey CF and Gimbrone MA (1993). Proc. Natl. Acad. Sci., USA, 90, 4951-4595.

De la Rosa EJ, Bondy CA, Hernandez-Sanchez C, Wu X, Zhou J, Lopez-Carranza A, Scavo LM, de Pablo F (1994). Eur J Neurosci, 6, 1801-1810.

Satokata I and Maas R (1994). Nature Genet., 6, 348-356.

Shiang R. Lidral AC, Ardinger HH, Buetow KH, Romitti PA, Munger RG, Murray JC (1993). Am. J. Hum. Genet. 53: 836-843.

Shields ED, Bixler D, Fogh-Andersen P (1981). Clin. Genet., 20: 13-24.

Shields ED, Bixler D, Fogh-Andersen P (1981). Cleft Palate J. 16: 1-6, 1979.

Spielman RS, McGinnis RE and Ewens WJ (1993). Amer. J. Hum. Genet., 52, 506-516.

The Treacher-Collins Syndrome Collaborative Group (1996). Nature Genet., 12, 130-136.

Werler MM, Lammer EJ, Rosenberg L and Mitchell AA (1991). American Journal of Epidemiology, 134, 691-698.

Williams A, Shaw WC and Devlin HB (1994). British Medical Journal, 309, 1552.

P32039wo